Statistical Considerations of Food Allergy Prevention Studies

Henry T. Bahnson, MPH[a], George du Toit, MB, BCh, FRCPCH[b,c], and Gideon Lack, MB, BCh, FRCPCH[b,c] *Seattle, Wash; and London, United Kingdom*

Clinical studies to prevent the development of food allergy have recently helped reshape public policy recommendations on the early introduction of allergenic foods. These trials are also prompting new research, and it is therefore important to address the unique design and analysis challenges of prevention trials. We highlight statistical concepts and give recommendations that clinical researchers may wish to adopt when designing future study protocols and analysis plans for prevention studies. Topics include selecting a study sample, addressing internal and external validity, improving statistical power, choosing alpha and beta, analysis innovations to address dilution effects, and analysis methods to deal with poor compliance, dropout, and missing data. **(J Allergy Clin Immunol Pract 2017;5:274-82)**

[a]Immune Tolerance Network, Benaroya Research Institute, Seattle, Wash
[b]Department of Pediatric Allergy, Division of Asthma, Allergy and Lung Biology, King's College London, London, United Kingdom
[c]Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom
Open access for this article was funded by King's College London.
Conflicts of interest: H. T. Bahnson has received research support from the National Institute of Allergy and Infectious Diseases (NIAID)/National Institutes of Health (contract nos. HHSN272200800029C and UM2AI117870). G. du Toit has received research support from the NIAID (NO1-AI-15416 [contract] and UM1AI109565 [grant], covering salary) and the UK Food Standards Agency; has received a contribution to NIAID contract/grant from the Food Allergy Research & Education; has received a contribution to KCL Division of Asthma Allergy & Lung Biology from Medical Research Council (MRC) & Asthma UK Centre; has received the Biomedical Research Centre (BRC) award to Guy's and St Thomas' National Health Service (NHS) Foundation from the UK Department of Health through the National Institute for Health Research (NIHR); and has received support for pediatric allergy clinical trial's unit from the National Peanut Board. G. Lack has received research support from the NIAID (NO1-AI-15416 [contract] and UM1AI109565 [grant]), and UK Food Standards Agency; is on the DBV Technologies scientific advisory board; has received a contribution to NIAID contract/grant from Food Allergy Research and Education; has received a contribution to KCL Division of Asthma Allergy & Lung Biology from MRC & Asthma UK Centre; has received the BRC award to Guy's and St Thomas' NHS Foundation from the UK Department of Health through NIHR; has received support for pediatric allergy clinical trial's unit from the National Peanut Board; has received discounted Bamba peanut snack from Osem; and has stock/stock options in DBV Technologies.
Received for publication October 6, 2016; revised manuscript received and accepted for publication December 20, 2016.
Corresponding author: Gideon Lack, MB, BCh, FRCPCH, Children's Allergy Unit, 2nd Fl, Stairwell B, South Wing, Guy's and St Thomas' NHS Foundation Trust, Westminster Bridge Rd, London SE1 7EH, UK. E-mail: Gideon.Lack@kcl.ac.uk.
2213-2198

http://dx.doi.org/10.1016/j.jaip.2016.12.007

Key words: *Prevention studies; Food allergy; Statistical considerations; Complier average causal effect; LEAP; EAT; Dilution effects; Dropout; Imputation; Missing data; Tipping point analysis; Type I error; Type II error*

The prevalence of food allergy has been on the rise over the last 30 years with 6% to 8% of children being affected worldwide.[1,2] Currently, there is no cure for IgE-mediated food allergy and the main treatment remains avoidance; thus, understanding the cause and developing strategies for the prevention of allergy has been at the forefront of current allergy research. The past decade has seen an increase in trials aimed at the prevention of food allergy through early life nutritional interventions.[3-6] These prevention trials, in contrast to therapeutic trials, apply to subjects at risk of developing a future food allergy and therefore tend to be drawn from an at-risk pediatric population.

Although prevention trials can lead to valuable public health recommendations (eg, childhood vaccination or early consumption of peanuts[7]), their design, implementation, and interpretation pose unique and significant challenges. Prevention trials often take longer to complete, show smaller treatment effects, and require larger numbers of participants than do studies designed to test a therapy on a preexisting illness. Because participants are ostensibly healthy, the risk-benefit ratio of aggressive intervention is often shifted toward safer, more conservative strategies. Conservative interventions can lead to smaller treatment effects and therefore require larger sample sizes to achieve adequate power. Moreover, a drug's side effects are experienced by only the small number of people treated with the drug. Conversely, harmful effects resulting from a broadly applied public policy recommendation can eliminate the public utility of the intervention because adverse events will be experienced over a large portion of the population.

The data analysis of prevention studies can also present unique challenges. Because a large proportion of the study sample typically does not develop the disease of interest, these participants can dilute or add variability to the metrics used to evaluate safety and efficacy. Prevention trials are often longer in duration to coincide with the incidence of disease. Unfortunately, participants enrolled in lengthy studies tend to have higher rates of dropout and lower rates of compliance, especially if they perceive little or no immediate benefit.[8] Nevertheless, the many challenges that exist with conducting and analyzing prevention trials can be addressed with appropriate study design features and statistical methodologies.

This article focuses on the statistical challenges of food allergy prevention studies. However, the concepts apply to the design and analysis of nearly all prevention trials and particularly to diseases with low prevalence. Examples are drawn from 2 recently published randomized controlled prevention trials: Learning Early About Peanuts (LEAP) and Enquiring About Tolerance (EAT). Briefly, the LEAP and EAT studies enrolled 640 and

Abbreviations used
CACE- complier average causal effect
EAT- Enquiring About Tolerance
ITT- intention-to-treat
LEAP- Learning Early About Peanut Allergy

1303 infants, and took 7 and 7.5 years to complete, respectively. The LEAP study participants were recruited from an at-risk population (severe eczema and/or egg allergy) and the EAT study participants were recruited from a general population of exclusively breast-fed infants. At completion of the trials, the peanut allergy prevalence in the control group was 17.2% in the LEAP study and 2.5% in the EAT study, and compliance with the intervention was 92% and 54%, respectively. Using these trials as the main examples, topics in the following areas of food allergy prevention studies will be addressed:

1. Study Design: enrollment criteria, external validity
2. Statistical Power: choice of alpha and beta and 1- or 2-sided hypothesis testing
3. Analysis innovations to address dilution effects
4. Analysis methods to deal with poor compliance, dropout, and missing data

STUDY DESIGN

Whom to enroll?

Determining whom to enroll for a prevention study involves additional challenges not typically present for a therapeutic trial. When testing a new drug or therapy, participants with the disease of interest need to be identified and enrolled. Conversely, in prevention trials, participants must be enrolled before the illness presents. If disease prevalence is low (eg, peanut allergy at ~2%), a random sample from the overall population needs to be very large to provide sufficient power. Moreover, the large proportion of participants unaffected by an illness often perceives less immediate study benefit. Thus, they may be unmotivated to comply with an intervention and continue study participation. Poor compliance and dropout impair analysis interpretation by decreasing statistical power and producing results that lack internal or external validity. Therefore, selecting a high-risk population can offer key advantages. Figure 1 illustrates a simulation study in which the intervention effect (80%) and sample size (n = 1000) are held constant. The selection criteria are made more restrictive to enrich the study sample with a higher proportion of, for example, peanut allergy. The analysis demonstrates vastly lower *P* values with more restrictive enrollment criteria. This same concept also applies to subgroup and covariate-adjusted analyses, which can be specified using baseline factors known to be associated with the outcome of interest. These subgroups and covariates, if specified *a priori*, can form more powerful primary analysis comparisons within a larger, population-based sample.

Conversely, a study that is too restrictive in its selection criteria can lack external validity if the participants poorly represent the general population. A method to address this shortfall is to sample from the overall population using factors (eg, eczema severity) known to be associated with the development of food allergy. If the resulting sample has a wide distribution of the factor and the approximate distribution is known in the larger population, these prevalence estimates can be used to back-calculate the intervention

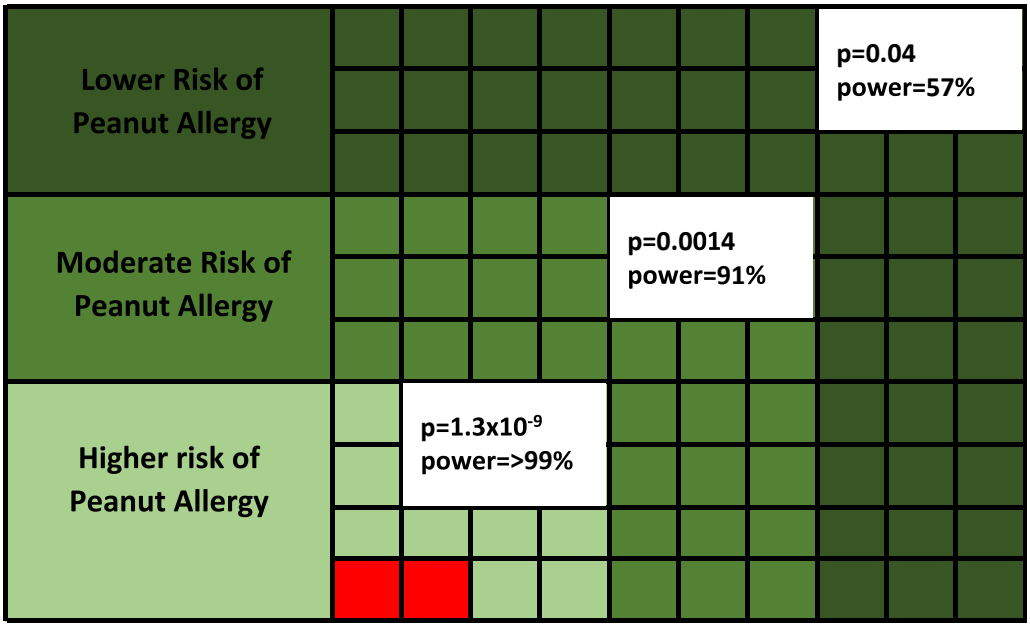

FIGURE 1. Three sampling strategies (different shades of green) are shown from an overall population with a prevalence of peanut allergy of 2% (depicted by the 2 red squares out of a 100 green squares). The outer band represents a population-based study where all squares are randomly sampled to produce a representative sample with a 2% prevalence of peanut allergy. As the bands move inward, the selection criteria is more restrictive and the proportion with peanut allergy increases from 2% to 4% to 12.5%. The annotated P values and power levels (white blocks) are from Fisher exact tests between the simulated control and treatment groups using a 2-sided test of significance at alpha = 0.05. Table I and Figure 1 demonstrate a dramatic increase in statistical power with more focused selection criteria, despite the intervention effect (80%) and sample size (n = 1000) being held constant in each simulation.